GB03/000241

The Patent Office

INVESTOR IN PEOPLE

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ

PRIORITY DOCUMENT

SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

REC'D 22 APR 2003

WIPO PCT

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

R. Mahoney

Signed

Dated 18 March 2003

24JAN02 E690268-2 D00027
P01/770[illegible]0-0201523.8

P[illegible]nts Form 1/77

THE PATENT OFFICE
K
23 JAN 2002

Patents Act 1977
(Rule 16)

The Patent Office

1/77

The Patent Office
Cardiff Road
Newport
Gwent NP9 1RH

Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

1.	**Your reference**	59.76152
2.	**Patent application number** *(The Patent Office will fill in this part)*	**0201523.8** 23 JAN 2002
3.	**Full name, address and postcode of the or of each applicant** *(underline all surnames)*	Mohamed Raafat El-Gewely Uranusveien 38, 9024 Tomasjord Norway
	Patents ADP number *(if you know it)*	8310179001
	If the applicant is a corporate body, give country/state of incorporation	
4.	**Title of the invention**	Molecular Libraries
5.	**Name of your agent** *(if you have one)*	Frank B. Dehn & Co.
	"Address for service" in the United Kingdom to which all correspondence should be sent *(including the postcode)*	179 Queen Victoria Street London EC4V 4EL
	Patents ADP number *(if you know it)*	166001

		Country	Priority application number *(if you know it)*	Date of filing *(day / month / year)*
6.	**If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and** *(if you know it)* **the or each application number**			

		Number of earlier application	Date of filing *(day / month / year)*
7.	**If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application**		

8.	**Is a statement of inventorship and of right to grant of a patent required in support of this request?** *(Answer 'Yes' if:* *a) any applicant named in part 3 is not an inventor, or* *b) there is an inventor who is not named as an applicant, or* *c) any named applicant is a corporate body.* *See note (d))*	No

Patents Form 1/77

9. Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document	
Continuation sheets of this form	0
Description	43
Claim(s)	0
Abstract	0
Drawing(s)	7 + 7
10. If you are also filing any of the following, state how many against each item.	
Priority documents	
Translations of priority documents	
Statement of inventorship and right to grant of a patent *(Patents Form 7/77)*	
Request for preliminary examination and search *(Patents Form 9/77)*	
Request for substantive examination *(Patents Form 10/77)*	
Any other documents *(please specify)*	
11.	I/We request the grant of a patent on the basis of this application. Signature Date 23 January 2002
12. Name and daytime telephone number of person to contact in the United Kingdom	Rebecca Gardner 01273 244200

Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

a) If you need help to fill in this form or you have any questions, please contact the Patent Office on 0645 500505.
b) Write your answers in capital letters using black ink or you may type them.
c) If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s) of the form. Any continuation sheet should be attached to this form.
d) If you have answered 'Yes', Patents Form 7/77 will need to be filed.
e) Once you have filled in the form you must remember to sign and date it.
f) For details of the fee and ways to pay please contact the Patent Office.

76152.358

Molecular Libraries

The present invention relates to molecular libraries in particular to molecular libraries which can be screened to identify members of the library with certain characteristics.

The complex interactions and pathways between biomolecules in living organisms are the focus for much research, particularly as scientists strive to develop new candidate drugs for the diagnosis, prevention or treatment of disease. Of particular interest are proteins which have functionally active surfaces and pockets which are characterised by the three dimensional shape of the protein molecule and particular reactive amino acid side groups. These proteins are able to interact, associate or bind with a variety of other biomolecules, including other proteins, nucleic acids, organic molecules, ions etc.

There is a great desire to identify molecules which can undergo specific interactions with target molecules, e.g. proteins, such molecules may be able to act as agonists or antagonists of enzyme or receptor activity for example, and more recently it has been proposed that the three dimensional shape of target proteins may be altered through interactions with small peptides (Nature Biotechnology letter El-Gewely 1999, Vol 17, p210). In this way the wild-type function of a mutant protein can be restored.

A variety of molecular libraries which can be screened to identify molecules with desired properties are known in the art. These may comprise organically synthesised chemicals or chemicals produced by combinational chemistry. It is expensive to generate very large libraries of such organic chemicals even using combinatorial techniques and so interest has

turned to biological systems.

The best known of these is phage display where peptides are expressed on the surface of a filamentous phage as a fusion protein with a coat protein. Ideally each phage will carry only one such peptide (although it will generally have several copies of that molecule) on its surface and these are then screened *in vitro* for their ability to bind to a target ligand. The phage provide a useful link between genotype and phenotype, as the phage which has a useful ligand on its surface will also have the nucleic acid inside it which codes for that peptide ligand. The phage can be cultured, the nucleic acid amplified and then sequenced to determine the identity of the ligand exhibiting desirable properties. However, only ligands in the form of fusion proteins with coat proteins can be investigated according to this method and it is only possible to perform a rather limited *in vitro* binding assay. Much more useful information could be generated if it was possible to analyse the performance of the members of the molecular library *in vivo*, in an environment similar to that which would be found inside the cells where the molecule is intended to act.

Another biologically based system which provides a molecular library is known as Aptamers/SELEX (Ellington and Conrad 1995, Gold et al. 1997 etc. SELEX stands for Systematic Evolution of Ligands by Exponential Enrichment. According to this technique, single or double stranded nucleic acid molecules that bind to other molecules are investigated, e.g. in order to find a small molecule - in this case a nucleic acid molecule, that binds to a target protein (e.g. a vitamin, amino acid, antibiotic, peptide etc.). The short nucleic acid molecules with strong and specific binding capabilities are the "aptamers". Again, these systems only allow *in vitro* screening of molecules.

Thus, there remains a need for a relatively cheap

molecular library whose members can be analysed *in vivo* without the need to generate fusion proteins which as a whole may not reflect the potential binding activity of the individual library molecules themselves. These problems have been addressed by the present invention, which according to one aspect provides a library of nucleic acid constructs which can express free peptides in an intra-cellular environment, said peptides having the sequence M-G/M/V-$(X)_n$, wherein n is an integer from 3 to 18, M is methionine, G is glycine, V is valine and each X, which may be the same or different, is any genetically encoded amino acid.

The above defined free peptides are optionally expressed as a fusion peptide with a cleavable signal sequence responsible for localization or translocation of the peptide. Such signal sequences are well known in the art and will typically be no more than 30, preferably no more than 20 amino acids in length.

The integer represented by the letter n in the above formula is between 3 and 18, preferably between 3 and 10 and typically between 3 and 5.

Thus in each of the expressed peptides the first residue is methionine, the second residue is glycine, methionine or valine and the third and subsequent amino acids may be any of the genetically coded amino acids. The universally accepted single letter code system for amino acids is used throughout.

It is well known that there are 20 genetically coded amino acids and any of these may be present in the encoded peptides. There will be between 3 and 18 X residues which may be the same or different and it is envisaged that for a given library, once the desired length of the peptides has been set, representatives of most if not all available combinations of residues will be present. Thus for example, using the single letter code for amino acids, where the desired peptides are pentamers, the library would encode most or all of the

following peptides, which are themselves only listed by way of example, MMACD, MMACE, MMACF, MMACG, MMADE, MMADF, MMADG, MGACD, MGACE, MGACF, MGACG etc. Preferably the second amino acid is glycine.

A library of constructs according to the invention may encode peptides of different lengths but typically a library will encode only peptides of 1-3 different lengths, often of only one length. Thus a library will typically have constructs which encode just pentamers or just hexamers etc.

The present invention provides a library of constructs which in turn is capable of generating a library of peptides which can be screened for their activity in a given intra-cellular environment. Thus, according to the present invention, the library of molecules which is actually investigated/screened is made up of peptides, the library as a whole providing a repertoire of different shapes. The "shapes" will also have regions of positive and negative charge and various reactive groups which will impact on their *in vivo* activity.

In a further aspect the present invention provides a library of peptides, each member of said library having the sequence M-G/M/V-$(X)_n$, wherein n is an integer from 3 to 18, M is methionine, G is glycine, V is valine and each X, which may be the same or different, is any genetically encoded amino acid.

The libraries of the present invention have at least 50, typically at least 100, preferably at least 500 members more preferably at least 2000. The different constructs are recognisably part of a library as they will generally differ only in the region which encodes the peptides of formula M-G/M/V-$(X)_n$ defined above.

The term "nucleic acid construct" refers to a nucleic acid molecule which contains a coding region for the peptides as defined herein operably linked to a

promoter region. The constructs are typically in the form of expression, e.g. plasmid vectors and other expression vectors such as phagemid or lambda expression vectors could also be used. The constructs are designed to allow expression in prokaryotic or eukaryotic host cells and thus a mammalian, bacterial or yeast etc. expression vector is be selected as appropriate. Suitable expression vectors for use with different cell types are well known in the art. The library of peptide molecules encoded by the library of constructs of the invention can be analysed for their ability to interact with a given target system or protein. If the target protein is mammalian it follows that the library will ideally be expressed in mammalian cells and the expression vector will be selected accordingly.

Thus the invention provides libraries for use in any host cell for which a suitable promoter to drive expression is known. Modifications to nucleic acid libraries disclosed herein which would be necessary for expression in a different host cell system are within the competence of the skilled man, so a range of bacterial, eukaryotic or yeast compatible libraries may readily be developed. Interestingly the Shine-Dalgarno sequence which is required for expression in bacteria does not hinder expression in eukaryotic cells.

It should be understood that the construct library is capable of generating a repertoire of different molecules with different three dimensional configurations and different functional groups. As with a chemical library, this same library of peptides can be used in many different screening methods; whether it be for allosteric inhibition of an enzyme, the ability to competitively inhibit the activity of an intra-cellular messenger at a receptor binding site, to control nucleic acid transcription or translation, or to restore the function of a mutant protein through interaction therewith which alters the mutant protein's three

dimensional shape.

Whereas a phage display library is used in an in vitro binding assay, the libraries of the present invention, because the peptides are expressed as free peptides inside the cell, may be used in an *in vivo* screening method to investigate the activities of the peptide ligands in normal cell conditions. Thus toxic molecules may be excluded at an early stage in testing.

The cells in which the library of peptide molecules are to be expressed can internalise the expression vectors into which the ligand repertoire has been constructed and there is thus no need to either synthesize peptides or purify peptide libraries. Preferably, the expression vectors are capable of autonomous replication to prevent the dilution of transfected ligand-encoding plasmids during cell division or the integration of the (e.g. plasmid) vector into the host cell chromosome, which causes difficulties in sequence analysis of the region of the nucleic acid within a cell which encodes the peptide ligand of interest.

Recombinant technology to generate expression vectors is well known and a suitable method for generating the library of constructs incorporating the desired sequence variation is described in the Examples. The variation in the nucleic acid sequence which translates into the different peptide ligands is typically generated through the synthesis of a master primer which has regions of known sequence flanking the region which encodes the peptide of formula $M\text{-}G/M/V\text{-}(X)_n$ defined above. As the first amino acid is always methionine, the first three nucleotides will be ATG, in the preferred case where the second residue is glycine, the next three nucleotides will be GGA and the remaining triplets follow the formula NNK, wherein N is A, C, T or G and K is G or T. This can allow for any nucleotide in any of the Y positions. The primer is generated adding

particular nucleotides in turn until the NNK regions are reached when all 4 nucleotides are added in equal molar amounts for the addition of 2 nucleotides to the growing strand and then G and T are added in equal molar amounts for the third position in the triplet. This cycle of additions is repeated until nucleic acid encoding for the peptides of the required length is generated. There are many commercially available machines or services for synthesis of oligonucleotides, e.g. as provided by Sigma and Life Technologies.

In a further aspect, the present invention provides a method of identifying an active peptide ligand of formula M-G/M/V-$(X)_n$ as defined above which comprises:

a) transforming a host cell population with a library of nucleic acid constructs as defined herein;

b) culturing the transformed host cells under conditions suitable for intra-cellular expression of the free peptides of formula M-G/M/V-$(X)_n$;

c) analysing the host cell population to determine the effect of the expressed free peptides on a reporter system; and

d) identifying the peptide of formula M-G/M/V-$(X)_n$ in those cells in which the reporter system indicates a positive response.

Suitable transformation and culturing techniques are well known in the art. Suitable reporter systems are discussed in detail below and will depend on the desired activity of the ligands. The reporter system may be directly or indirectly related to the desired ligand function. For example, if the ability of a ligand to activate transcription or translation of a gene conferring antibiotic resistance is being investigated, culturing the host cells in the presence of that antibiotic would readily identify successful ligands as cells which expressed non-successful ligands would not grow. Given the size of the peptides, activation is unlikely to be through direct nucleic acid

binding but through interaction with the protein machinery which performs and/or controls transcription/translation.

Thus analysis of the host cell population may simply be by observation of growing plaques or through detecting a particular colour if the β-gal reporter system is used. In step d) the active peptide is typically identified by analysis of the nucleic acid encoding it.

~~A library of molecular shapes, i.e. ligands,~~ for *in vivo* screening has a wide variety of uses. Of particular interest is the potential for members of a library to encode peptides which can take part in protein therapy through protein function restoration. It has recently been shown that while a mutant protein may have an altered three-dimensional structure due to a non-wild type primary structure, the three-dimensional structure can be corrected without changing the primary structure. Small molecules, e.g. peptides, may be used which interact with the mutant protein to modify its three dimensional structure and thus its activity. Proteins are known to change their conformation on binding or associating with a range of different classes of molecules, e.g. other proteins, nucleic acids, prosthetic groups such as heme, ions such as Ca^{2+} etc.

Peptide ligands identified according to the above described method can be used therapeutically to alter the three-dimensional structure of a mutant protein and thus restore or alter the function or activity of that mutant protein. Typically the ligand can restore, in whole or in part, wild type protein function. Peptide ligands identified according to the above screening method may be chemically modified before being used in therapy.

It has been found that very small peptides can restore or modify the function of target proteins through interacting with them *in vivo* and causing

changes in their conformation. The free peptides expressed by the library of nucleic acid constructs will preferably have only 3-8 amino acids, particularly preferably 3-6 amino acids. If a signal sequence is used, these preferred lengths refer to the peptides generated after cleavage of the signal sequence.

In a further aspect, the present invention therefore provides for the use of a library of nucleic acid constructs as defined herein in a method of screening a library of molecules for the ability of members of that library to restore or modify the function of a target protein in an intra-cellular environment, which method comprises introducing the library into cells which have a reporter system which allows the identification of those cells in which the function of the target protein has been restored or modified.

Such a screening method gives information about individual members of the library of molecules. The method is performed in an intra-cellular environment (i.e. *in vivo*), and therefore all toxic compounds will be weeded out very early on. This reduces the cost of screening and of pre-clinical and even phase I clinical trials. The method is able to indicate in a qualitative possibly also quantitative manner the performance of individual molecules in the test system. A target protein and suitable host cells are selected and the method designed so that the reporter system is able to give information about the ability of each molecule to restore or modify the function of the target protein.

The reporter system is typically designed to measure restoration of a particular function of the target protein, for example the ability to act as a DNA binding transcription factor. Thus restoration of function refers to a return to wild-type function of the protein in at least one physiologically relevant and measurable respect.

Suitable target proteins for restoration of function will include those whose presence in mutated form is associated with a disease state. The mutated version of the protein will typically have an altered 3-dimensional structure which affects its ability to interact with other molecules (ions, intra-cellular organelles and other cell components etc.). The members of the molecular library will be screened for their ability to interact with the mutant target and thus alter 3-dimensional structure. The 3-dimensional structure *in vivo* may be closer in one or more respects to the wild-type 3-dimensional structure as a result of successful interaction with a member of the library but protein function may be restored through a further compensating change in 3-dimensional structure which has a return to wild-type function as a 'net' result. Preferred target proteins will be those wherein point mutation(s) are the cause of disease. Particularly preferred are transcription factors (DNA binding proteins) such as p53, but others would be enzymes, peptide hormones or receptor molecules. All diseases caused by monogenic Mendelian mutations could be targets for treatment with molecules identified according to the screening methods of the present invention. These include genetic (i.e. hereditary) diseases, cancer and symptoms of aging caused by mutations. The host cells should provide a suitable model for determining whether the members of the library are capable of causing the desired change in target protein function with a view, typically, to therapeutic administration of successful members of the library or a derivative thereof. Thus, where protein function restoration is required, the host cells will typically be derived from cancerous or other diseased cells and naturally produce the dysfunctional target protein. Suitable cell lines are available, for example in cell culture collections such as the ATCC and may be derived from osteosarcoma, adenocarcinoma etc.,

or indeed any established cell line that expresses the mutant protein. Cell lines that lack any expression of the protein of interest can also be used after introducing and expressing the gene encoding the mutated protein.

Standard transfection techniques may be used to introduce the nucleic acid constructs into the host cells in which expression of the free peptide ligands is to take place.

The appropriate reporter system for a given screening method will depend upon the nature of the target protein under investigation. The reporter system is chosen to be responsive to the target protein and thus be capable of indicating the functional status of the target protein. The mammalian p53 protein illustrates the principle behind a suitable reporter system. P53 functions in mammalian cells mainly through the ability to act as a transcription factor and is a transcriptional activator of several genes associated with cell cycle regulation such as p21, Bax, CD95(Fas/Apo-1) and 4-3-3σ (HME1). It can also down regulate the transcription of cell-cycle-regulating genes such as Bcl2, Cdc2 and cyclin. P53 acts as a checkpoint, monitoring DNA damage and regulating cell cycle progression. Loss of p53 activity predisposes cells to the acquisition of oncogenic mutations and may favour genetic instability - 90% of mutations reported in the p53 mutation database are found in the DNA binding domain. Mutations in p53 can lead to cancer formation where p53-mediated apoptosis is deficient due to the mutation.

P21 is one of the p53-transactivated genes that are critical in cell cycle control and many p53 mutants found in cancer cells lose the ability to transactivate p21 transcription. Thus the transactivation level of p21 promoter in a cell reflects whether the p53 protein expressed in that cell has wild type function. As

discussed in more detail in the Examples, a p53 reporter system was constructed by cloning human p21 promoter upstream of the puromycin resistance gene. The reporter construct can provide human cells with puromycin resistance only in the presence of wild type or wild-type functioning p53 that transactivates the p21 promoter. Such a reporter system offers significant practical benefits as a life-death selection system is possible where only cells with wild-type-functioning p53 can survive when grown in the presence of puromycin.

Target proteins are preferably DNA binding proteins which up or down regulate the expression of other genes through binding to promoter or enhancer regions. Their native gene targets may be utilised to report on whether a target protein function has been restored or modified, but preferably their target DNA binding regions will be operably linked to reporter genes such as genes conferring antibiotic resistance, or which encode fluorescent proteins such as GFP (green fluorescent protein) etc.

A library of nucleic acid constructs according to the present invention is conveniently referred to herein as a "Microgenex" library and the stretches of nucleic acid which encode the peptide ligands of interest as "Microgenes".

The invention will now be described in more detail in the following non-limiting Examples and with reference to the figures in which:

<u>Figure 1</u> shows in schematic form the general strategy for construction of a library of nucleic acid constructs according to the present invention.

<u>Figure 2</u> illustrates the strategy of primer design for construction of a Microgenex library. Restriction enzyme sites within primers as well as the encoded amino acids are shown

below the primer sequences. M = methionine, G = glycine and X = any of the 20 genetically coded amino acids.

<u>Figure 3</u> illustrates the construction of a peptide library for expression in mammalian cells.

<u>Figure 4</u> illustrates the construction of a peptide library for expression in bacteria *(E. coli)*.

<u>Figure 5</u> gives a map of the constructed Microgenex library of the Examples that can be used for expression in mammalian cells.

<u>Figure 6</u> gives a map of the constructed Microgenex library of the Examples that can be used for expression in *E. coli*."Cm-R" refers to chloramphenicol resistance gene. "p15A" is replication origin compatible with several other replication origins used in molecular biology and biotechnology. P_{BAD} promoter is suppressed by araC but can be induced by L-arabinose. Peptide library coding sequence was inserted under the P_{BAD} promoter between KpnI and HindIII sites. SD sequence is a ribosome-binding site to help in the translation of the peptides in *E. coli*. The first two amino acids at the N-terminal ("M" for Met, and "G" for Gly) can stabilize the peptides inside cells. The two stop codons assure the termination of translation while the terminator ("term") sequence will terminate the transcription driven by the P_{BAD} promoter. The DNA fragment of the peptide library was generated by PCR, using Master, Right and Left primers as described in the Examples.

Figure 7 provides a schematic representation of the generated peptide library plasmid. Each peptide in the library is composed of three random amino acids after two N-terminal amino acids Methioline (Met, or M) and Glycin (Gly, or G). The next three `X' stands for any amino acid that forms a library consisting of randomly distributed three-amino-acid peptides. The peptide-coding DNA-fragment was inserted into the mammalian expression vector pCEP4 between its KpnI and HindIII site under the CMV promoter. The plasmid has OriP that enables the replication of the plasmid inside mammalian cells. Selection marker expression cassette in the plasmid provides the plasmid-transfected cells with hygromycin resistance.

Examples

A: Strategy for Microgenex construction:

Materials and Methods:

1- Primers.

All primers used to construct Microgenex libraries are listed in Table 1.

Table 1. Primers: Primers used to construct Microgenex libraries.

No	Name	Sequence	bases
1	Master	AAGAGCTCGGTACCAAGAAGGAGTTTACATATGGGANNKN NKNNKTGATAAGGATCCAAGCTTGAATTCAG	71
2	Left	AAGAGCTCGGTACCAAGAAGGAG	23
3	Right	CTGAATTCAAGCTTGGATCCTTATC	25

Primers used for DNA sequence are listed in Table 2.

Table 2. The Sequencing primers used for the Microgenex verifications.

No	Name	Sequence	bases
1	pCEP-f	AGAGCTCGTTTAGTGAACCG	20
2	pCEP-r	GTGGTTTGTCCAAACTCATC	20
3	Left*	AAGAGCTCGGTACCAAGAAGGAG	23
4	Right*	CTGAATTCAAGCTTGGATCCTTATC	25

*These primers can also be used to verify sequence of any Microgenex library. Other primers within the used cloning vector can also be used.

2- Cloning Vectors/ plasmids

pCEP4 (Invitrogen): for mammalian expression.

pBAD33 (Dr. Beckwith, Harvard; Guzman et al., 1995 J. Bacteriol. 177(14) 4121-4130) for *Escherichia coli* expression.

3- *Escherichia coli* strain used for Transforming Constructed Microgenex libraries:

The Electrocompetent DH10B™ (Life Technologies, GibcoBRL) was used to transform the ligated Microgenex expression libraries.

4- PCR conditions for Microgenex amplifications

4.1. Concentrations of primers:

Master primer 25 PM
Right primer 100 PM
Left primer 100 PM

4.2. PCR reactions:

At least two different Polymerases gave good results: Taq polymerase (Life Technologies, GibcoBRL) and rTth DNA polymerase (Perkin Elmer).

4.2.1. Standard Taq polymerase as for example in SuperMix (Life Technologies, GibcoBRL).

Prepare 1 to 10 PCR tubes with the following:

90 μl SuperMix
2 μl Master Primer (25 PM)
4 μl Right Primer (100 PM)
1 Left Primer (100 PM)

4.2.2 The XL PCR, extra long (rTth DNA polymerase) (Perkin Elmer)

Also this PCR polymerase worked very well after adding buffers according to the manufacturer, but using primers as above and also the PCR cycles as indicated

below.

4.2.3. PCR Cycles

Best Conditions were found by making a hot start at 95 for 3-5 minutes. PCR cycles were preformed only using 25 cycles. Example for PCR Cycle:

Hot start	95 °C	3 min.
25 cycles		
Denaturation	94 °C	1 min.
Aneaing	38 °C	1 min.
Elongation	60 °C	1 min.
Hold at	4 °C	

5. Ligation of PCR products to the corresponding expression vector:

PCR product was first treated with phenol-chloroform, chloroform and then precipitated with 3 volumes of cold ethanol. The DNA pellet was washed with cold 70% ethanol, dried and dissolved in appropriate volume of H_2O. DNA was then restricted by KpnI and HindIII. Restricted DNA was gel-purified and subsequently ligated to the corresponding expression vector that has been also restricted with by KpnI and HindIII. Ligation mix was precipitated and resuspended in H_2O and subsequently used to transform, by electroporation, of the *Escherichia coli* strain DH10B™ (Life Technologies, GibcoBRL). Plating was done on LB ampicilin plates (200 μg/ml) for pCEP4, or chloramphenicol (40 μg/ml) for pBAD33.

6. Pooling of colonies:

Colonies from all the plates were pooled together, washed

with LB with ampicilin (200μg/ml) if the cloning vector contains beta-lactamase gene, such in the case of pCEP4. This step is important to get rid of cells that lack the plasmid. Pooled cell mixture was divided into small aliquots and stored at -70°C. Each aliquot contained millions of cells to ensure complexity of the Microgenex library.

7. DNA Isolation of Microgenex library:

One frozen portion can be used to inoculate a one-liter culture of LB containing the appropriate antibiotic. Standard plasmid DNA isolation protocols/kits are used to isolate the DNA suitable to transform, for example mammalian cells for ligand screening.

Master primer was used as a template for all constructed peptide expression libraries (Microgenex) whether the expression in mammalian cells or *Escherichia coli*. It could also be used for yeast or insect cell expression. Left and Right oligoes were used as forward and reverse primers, in order to generate double-stranded DNA fragments encoding repertoire of peptide library by PCR. The first ATG in this fragment encodes Met as the translation initiation site of the peptides. Gly, the second amino acid encoding by the DNA fragment, assures the peptides expressed to be stable inside cells according to the N-end rule. The following three repeats of NNK code (N for A, C, G, and T in equal molar ratio, K for G and T in equal molar ratio) for all possible amino acids thus form the peptide library. Two stop codons, TGA and TAA, are added right after the last NNK to stop peptide translation. The 5'-end KpnI site and 3'-end HindIII allows the fragments to be cloned into pCEP4 vector in the correct direction under CMV promoter.

The peptide-coding DNA-fragment was inserted into the mammalian expression vector pCEP4 between its KpnI and HindIII site under the CMV promoter. The plasmid has the replication origin, OriP, which enables the replication of the plasmid inside mammalian cells. Selection marker expression cassette in the plasmid provides the plasmid-transfected cells with hygromycin resistance.

After ligation of the KpnI/HindIII restricted PCR product, and the KpnI/HindIII restricted pCEP4 or pBAD 33, the DNA were electroporated into DH10B™ (Life Technologies, GibcoBRL) cells. Transformants were plated on LB plates with ampicilin (200μg/ml) or Chloramphenicol (40 μg/ml) depending on the coloning vector used in the construction of the Microgenex library. All colonies from all the plates were pooled together, washed with LB with ampicilin (200μg/ml), only if the cloning vector contains beta-lactamase gene, such in the case of pCEP4. Pooled cell mixture was divided into small aliquots and stored at -70°C. Each frozen portion of the Microgenex library can be used to inoculate and start a 1-liter-culture (with the appropriate antibiotic, in order to prepare the plasmid DNA. Insertions were verified by PCR of the generated Microgenex with Right and Left oligoes. Also, the forward and reverse primers of pCEP4were used for verifications for the constructed mammalian Microgenex. Similar PCR reactions were preformed using plasmid DNA prepared from a small number of randomly-picked colonies in order to estimate the cloning efficiency and the ratio of peptide-encoding plasmids to empty vectors in any given Microgenex. Note that all generated peptides will start with M and G amino acids followed by three random amino acids (X, X, X) It should be noted that the number of the random amino acids (X) can be changed by changing the number of the codes (NNK).

Results

Colonies in the range of 25000-75,000 on the plates were commonly obtained ensuring complexity/diversity of the cloned Microgenex repertoire. They were collected into 10 - 12 portions.

Summary of Microgenex libraries are listed in Table 3.

Table 3. Plasmids used for the construction of microgenex libraries

Plamid Name	Promoter	Insertion	Selection marker	Replication origin	Reference.
pCEP4	CMV	No	*E. coli*: Amp, Mammalian: Hygromycin B	*E.coli*: ColE, Mammalian: OriP	Vector from Invitrogen
Mammalian microgenex library		Microgenex library with amino acid sequence as: M-G-X-X-X.			This work, Fig.2.3.
pBAD33	pBAD	No	Chloramphenicol	p15A	Vector from Beckwith; Guzman, et al.,1995
E. coli microgenex library		Microgenex library with amino acid sequence as: M-G-X-X-X.			This work.

Plasmid prepared from one of such frozen aliquots was examined by PCR. When mammalian Microgenex library DNA was used as a template in PCR reactions using pCEP4 forward and reverse primers showed that there were insertions with expected size. PCR results from the 6 randomly picked colonies from the original mammalian Microgenex library plates showed that 4 of the colonies had the right construction (showed correct size of insertion) while two did not.

A test for the presence of Microgene insertions in a constructed mammalian library by PCR from 6 randomly picked colinies indicated that 4 out of the 6 colonies have the right size of insertion. Also, PCR results from 6 randomly picked colonies from the original *E. coli* Microgenex library plates showed that 4 of the colonies had the right construction, in that they showed the correct size of insertion.

B: Demonstrating the generation and use of a microgene library in identifying a protein binding ligand of interest

Construction of p53 reporter working in mammalian cells

The dysfunction of mutant p53 in cancer cells can result from not only the inability in binding to its specific recognized consensus DNA sequence, but also the mal-localization inside cells or incompetent to interact correctly with other transcription co-factors. The sub-cellular localization mechanism and transcription machinery in mammalian cells is different from that in prokaryotic organisms. The loss of specific DNA binding ability may engage the post-translational modification processes that are totally different from that in prokaryotic organisms. Thus a p53 reporter for mammalian cells has more advantages than that for prokaryotic cells.

P53 protein functions in mammalian cells mainly through its transcriptional factor activity. P21 is one of the p53-transactivated genes that are critical in cell cycle control. P53 can induce cell cycle arrest in G1 phase via transactivation of p21 gene once the cells are under stresses such as DNA damaging. Lots of p53 mutants found in cancer cells lose the ability to transactivate p21 transcription. This leads to not only the vast development of cancer, but also the resistance to cancer chemotherapy and radiotherapy that aim to damage DNA in order to kill cancer cells. Thus the transactivation level of p21 promoter in a cell reflects whether the p53 protein expressed in that cell has wild type function or not.

We constructed our p53 reporter by cloning human p21 promoter upstream of the puromycin resistant gene. It has been reported that basal transcription level from p21 promoter is very low. The reporter can provide human cells puromycin resistance only in the presence of wild type or wild-type-functioning p53 protein that transactivates p21 promoter, while leaving the cells sensitive to puromycin with dysfunctional mutated p53 protein. The advantage to use puromycin resistant gene is to set up a life-death selection: only the cells with wild-type-functioning p53 protein can survive from puromycin selection.

Materials and Methods

Plasmid and vectors used in this Example are WWP-Luc (gift from Prof. Stanbridge E.J.), pPUR (Clontech), and pUC18 (Pharmacia). Enzymes used for cloning are EcoRI (Promega), PvuII (Promega), SalI (Promega), SphI (Promega), T4 DNA ligase (New England BioLab), Klenow (Promega), and SAP (Shrimp Alkaline Phosphotase). Cell lines for the validation of our reporter are U-2 OS cell

(ATCC Number: HTB-96, Homo sapiens (human), osteosarcoma, bone, wild type p53), Saos-2 cell (ATCC Number: HTB-85, *Homo sapiens* (human), osteosarcoma, bone, p53 null), and SW480.7 (ATCC number: CCL 228, *Homo sapiens* (human), colorectal adenocarcinoma, colon, p53 mutant: R273H/ P309S). Medium and other chemicals include McCoy's medium, DME medium, MEM, FCS, calcium phosphate transfection reagents (0.1xTE (pH8.0), 2xHBS (pH7.4), 2M $CaCl_2$, 15% glycerol in HBS), PBS, and puromycin (Sigma).

WWP-Luc plasmid was restricted by EcoRI and SalI. The 2366bp fragment was gel-purified (1% agarose in 1x TBE) by freeze-thaw method. Vector pUC18 was restricted by EcoRI and SalI, dephosphorylized by SAP, and purified by phenol-chloroform. Then the p21 promoter fragment was subcloned into pUC18 vector by T4 DNA ligase, generating a plasmid named as p21uc. All drug resistant markers in the constructed plasimds, are listed in Table 1. Genes involved in the resistance are also listed.

Table 1. Drug resistant markers in constructed plasmids and their genes.

Drug	Resistant gene
Puromycin	Puromycin-N-acetyl-transferase *(pac)*
Neomycin	Neomycin phosphotransferase gene
Hygromycin B	Hygromycin B phosphotransferase
Histidinol	Histidinol dehydrogenase (hisD)

P21 promoter was cut out from p21uc first by EcoRI, blunted with Klenow and dNTP, and then by SphI. The 2450bp-p21-promoter fragment was gel-purified using the same method as above. Vector pPUR was restricted fully by PvuII and SphI, dephosphorylized by SAP, and gel-purified by the same method, which removed SV40 early promoter upstream of puromycin resistant gene. Finally

the p21 promoter was cloned into pPUR vector upstream of puromycin resistant gene by T4 DNA ligase, generating the p53 reporter named as p21ur.

Verification of the construction was carried out by restriction of the generated plasmid with HindIII. The original vector pPUR has only one HindIII site thus gives one band of 4257bp after gel electrophoresis. This HindIII site was removed while constructing the reporter p21ur. Two new HindIII sites were introduced into p21ur with the p21 promoter thus can give two bands, one at 2.4kb and the other at 3.9kb.

Validation of the responsiveness of p21ur to p53 was carried out by transfection of p21ur to osteosarcoma cell line U-2 OS that expresses wild type p53 protein, Saos-2 that is p53 null, and colon cell line SW480.7 that expresses R273H and P309S double mutated p53 protein. One day before transfection, $3x10^5$ U-2 OS, 1.5 $x10^5$ Saos-2, and $3x10^5$ SW480.7 cells were seeded to each well of a 6-well plate. Transfection was carried out by the calcium phosphate method, with 25μg of p21ur for each well. A 2-minute glycerol-shock was applied 3 hours after DNA was added to the cells and the transfected cells were incubated in MEM/10%FCS at 37°C with 5% CO_2 overnight. The transfected cells were transferred to new 6-well plates at the concentration of $5x10^4$ cells per well on the next day, and incubated in growth medium (McCoy's medium/10%FCS for U-2 OS and DME medium/10% FCS for Saos-2 and SW480.7) with puromycin. The same puromycin treatment was also given to the non-transfected cells as a control. Refreshed the medium with puromycin every three-day and observed growth status of the cells.

The cloned vector was restricted by HindIII, which proved that there is the p21 promoter, a ~2.4kb fragment, inserted upstream of the puromycin resistant gene. DNA electrophoresis showed that the restriction gave two bonds with correct theoretical sizes at 2.4 and

3.9 Kb.

Transfection of the reporter p21ur into U-2 OS cell line provided puromycin resistance to the transfected cells (Table 2). Cells transfected with p21ur kept alive after 5 days of puromycin treatment at 0.5μg/ml, while the untransfected cells died out. The transfected cells could not survive from higher concentration of puromycin because of the low expression level of p53 protein. Neither transfected Saos-2 cells (null p53) or SW480.7 cells (mutant p53) survived after 5 days of puromycin treatment at the concentration of 0.5μg/ml or higher, which proved that the resistance of p21ur-transfected U-2 OS cell to puromycin requires the existence of wild type p53 that transactivates p21 promoter. Thus it validated the proper responsiveness of the reporter p21ur to p53 in human cells.

Table 2. Validation of the reporter p21ur - its response to wild type p53 protein.

Puromycin (μg/ml)	Number of cells attached to the bottom of the well after 5 days of puromycin treatment	
	Untransfected U-2 OS	Transfected U-2 OS
0	Growing well	12600
0.5	0	11500
1	0	<5000
2	0	<5000
3	0	<5000

<u>Construction of peptide library expression in mammalian cells</u>

We have constructed a microgene library to provide a series of molecular shapes for testing. This 'microgenex' approach enabled us to screen among the expressed peptide repretoire and identify some that

could correct mutated p53 and could restore all of its downstream activity.

Selection of the functional peptides can be carried out either *in vivo* or *in vitro* according to the peptide library. The advantage of *in vivo* selection over *in vitro* selection is that it allows the peptides to carry out their function in an environment similar or even exactly the same as where they will later be applied in therapy. This makes the peptides more applicable and simplifies the modification step for the selected peptides. Besides, it can often detect weak and transient interactions.

Expression of the peptide library allows *in vivo* screening for possible peptide ligands that can adjust mutant p53 protein back to its wild type function. Once the peptides were constructed into an expression vector, they were easily internalized by cells. With the help of positive-selection reporters, the possible ligands were also readily identified by the recovery of ligand-encoding plasmids from surviving cells. Moreover, according to this technique, there is no need to either synthesize peptides or purify peptides, which can be expensive and labour intensive. An un-integrated, autonomously replicated mammalian expression vector was chosen to prevent the dilution of transfected ligand-encoding plasmids during cell division, or the integration of the plasmids into the host chromosomes, which causes difficulties in rescuing ligand from surviving cells.

Materials and Methods

pCEP4 (Invitrogen):

TTGACGCAAA TGGGCGGTAG GCGTGTACGG TGGGAGGTCT ATATAAGCAG AGCTCGTTTA

CAAT TATA pCEP Forward

GTGAACCGTC AGATCTCTAG AAGCTGGGTA CCAGCTGCTA GCAAGCTTGC TAGCGGCCGC

primer KpnI HindIII

TCGAGGCCGG CAAGGCCGGA TCCAGACATG ATAAGATACA TTGATGAGTT TGGACAAACC

EBV Reverse primer

ACAACTAGAA

Oligo REGP-10:

5'-AAGAGCTCGG TACCAAGAAG GAGTTTACAT ATG GGA NNK NNK NNK TGA TAA
KpnI M G X X X stop stop
GGATCCAAG CTTGAATTCA G-3'
HindIII

Oligo REGP-11:

5'-AAGAGCTCGGTACCAAGAAGGAG-3'
KpnI

Oligo REGP-12:

5'-CTGAATTCAAGCTTGGATCCTTATC-3'
HindIII

pCEP Forward primer:

5'-AGAGCTCGTTTAGTGAACCG-3'

EBV Reverse primer:

5'-GTGGTTTGTCCAAACTCATC-3'

KpnI, HindIII, T4 DNA ligase, 10xT4 DNA ligase buffer

Oligo REGP-10 was used as a template, while oligos REGP-11 and REGP-12 were used as forward and reverse primers, to generate double-stranded DNA fragments encoding a degenerated peptide library by PCR. The first ATG in this fragment encodes Met as the translation initiation site of the peptides. Gly, the second amino acid encoding by the DNA fragment, assures the peptides expressed are stable inside cells according to the N-end rule. The following three repeats of NNK (N for A, C, G, and T in equal molar ratio, K for G and T in equal molar ratio) code for all possible amino acids and thus form the peptide library. Two stop codens, TGA and TAA, are added right after the last NNK to stop peptide translation. The 5'-end KpnI site

and 3'-end HindIII allows the fragments to be cloned into pCEP4 vector in the correct direction under CMV promoter.

After ligation of KpnI/HindIII restricted PCR product and KpnI/HindIII restricted pCEP4, the DNA molecules were electroporated into XL1-blue cells and plated on LB plates with ampicilin. All colonies from the plates were poured together, delivered into small portions and stored at -70°C. Each portion of the transformed cells can be inoculated and start a 1-liter-culture in order to prepared the plasmids by Qiagen Maxi-Prep kit. Insertions were verified by PCR of the generated plasmids with oligos REGPD-11 and REGPD-12. Same PCR reactions were also applied to plasmids prepared from 6 randomly picked colonies so as to estimate the cloning efficiency and the ratio of peptide-encoding plasmids to empty vectors in the population. Note that all generated peptides will start with M and G amino acids followed by three random amino acids.

Results

There were ~25000 colonies on the plates. They were collected into 12 portions. Plasmid prepared from one of such portions was examined by PCR.

Gel electrophoresis showed that there were insertions with correct size. PCR results from the 6 randomly picked colonies from the original library plates showed that 4 of the colonies had the right construction (showed correct size of insertion) while two did not. The structure of the generated peptide library plasmid is shown in Figure 7.

<u>Screening for ligands that can correct p53 mutants in SW480.7 cells</u>

SW480.7 is a human colon cancer cell line that carries

R273H and P309S double mutations in its p53 gene. R273H is one of the most frequently occurred p53 mutations that have been observed in various cancer tissues. It has been reported that the p53 mutant in SW480.7 loses its function in transactivating its downstream genes such as p21, Bax, 14-3-3σ, and CD95. Our purpose of screening the peptide library constructed above is to find out one or more peptide ligands that can restore wild-type p53 function without alteration of the mutant p53 gene in the cell and disturbing cells with wild type p53. The choice of a human cancer cell line provides the real human environment in which the selected peptides can be modified, degraded, and transported to subcellular organelles, and thus assures the reliability of the selected peptides to function in human body.

Materials and Methods

Human colon cancer cell line SW480.7 with p53 mutations R273H and P309S (ATCC Number: CCL-228, Homo sapiens (human), adenocarcinoma, colorectal; colon), DMEM medium, FCS, PBS, trypsin, puromycin (sigma), hygromycin (sigma), Peptide library (constructed by our laboratory described as above), pCEP4 vector (Invitrogen), calcium phosphate transfection reagents (0.1xTE (pH8.0), 2xHBS (pH7.4), 2M $CaCl_2$, 15% glycerol in HBS).

DNA extraction reagents (TBS buffer (0.8% NaCl, 0.02%KCl, 0.3%Tris, pH7.4), TE buffer (pH8.0), DNA extraction buffer (10mM TrisCl, 100mM EDTA, 20μg/ml RNase, 0.5% SDS, pH8.0), Proteinase K (18mg/ml), chloroform, phenol (Tris saturated), ethanol, 7.5M ammonium acetate (pH 7.4)), XL1-blue electroporation competent cell, pCEP Forward primer, EBV Reverse primer, dye-terminator thermocyclic DNA sequencing kit, ALF express DNA automatic sequencing system.

The peptide library was transfected into SW480.7 using the

calcium phosphate method described in the reporter constructing part. Meanwhile pCEP4 was also transfected into the SW480.7 cells as a negative control. On the second day after transfection, puromycin was added into the medium at the concentration of 0.5μg/ml. Then the transfected cells were washed with PBS and changed with fresh medium with 0.5μg/ml puromycin every three days. After all the negative control cells died, all the library-transfected cells that still attached to the wells were collected by scraping into ice-cold TBS buffer.

Rescue of library plasmids was carried out by the extraction of total DNA from the collected puromycin resistant cells according to Blin and Stafford's method. The extracted DNA was then directly electroporated into the competent cell XL1-blue and plated on LB plates with ampicillin. All the colonies were picked up for plasmid preparation. The plasmids were first screened by size using agarose gel electrophoresis to get rid of the reporters. The rest of the plasmids were sequenced by ALF express DNA automatic sequencing system (Pharmacia) using EBV Reverse sequencing primer (Invitrogen) to reveal the DNA sequence, hence amino acid sequence of the selected peptide ligands.

Results

No pCEP4-transfected cells survived in the well after 7 days of selection with puromycin, while there were still some living library-transfected cells sticking to the bottom of the well.

Electroporation of the DNA extracted from the library-transfected cells after 7 days of selection gave ~150 colonies. Among them, 50 of the colonies have the right size of library plasmid and were sequenced. One ligand peptide sequence was found. The DNA and amino acid sequences are:

DNA: 5'-GGTACCAAGAAGGAGTTTACATATG GGA TGG TGT ACT TGA TAA
KpnI
Amino acid: M G W C T stop stop

DNA: GGATCCAAGCTT-3'
HindIII

Validation of the selected peptide ligand

The purpose of this experiment is to confirm that the selected peptide ligand indeed restores the mutated p53 protein to the wild-type p53 function in SW480.7 cells. The selection of peptide ligands was based on puromycin resistance. There might be a chance that the cells got the resistance with, or even without the help of the selected ligand via mechanisms other than p53-transactivated p21 promoter. Thus we should make sure that the mutant p53 in SW480.7 cells functions as wild-type p53 only if the selected ligand is added. Confirmation/validation was achieved in two different ways:

1- We observed the function of the peptide ligand on cell cycle control of SW480.7 cells, as well as that of U-2 OS (WTp53) and Saos2 cells (no p53 copy). The selected peptide ligand can be further verified only if it can induce cell cycle arrest or apoptosis in SW480.7 cells but neither in U-2 OS nor in Saos2 cells, which means the specificity of the selected ligand to p53 mutation in SW480.7.

2- We also tested the transcription levels of genes that are only transactivated by wild-type p53. If the transcription levels of these genes are highly induced in ligand treated cells compared with the negative control (cells transfected only with the empty vector (pCEP4). The selected

ligand/ peptide was proved to be capable of altering the mutant p53 function in SW480.7 into wild-type function.

P53 (wild-type) regulates different cell responses via transactivating different downstream genes. For example p53 causes cell cycle arrest at G1 phase by transactivation of p21 gene. While p53 causes cell arrest at G2 phase by transactivation of CD95 gene. Apoptosis is due to p53 transactivation of 14-3-3σ and/or Bax gene. Thus we measured p21, CD95, and 14-3-3σ gene transcription levels by RT PCR, to reveal the capability of the selected peptide ligand in restoring the wild-type p53 transactivation of the downstream genes. The expression level of β-actin was measured as a control. β-Actin transcription level was used to normalize total mRNA amount between samples for comparison owing to its constitutive expression in all cells.

Materials and Methods

SW480.7, U-2 OS, Saos 2, DMEM, McCoy's medium, FCS, PBS, trypsin, trypan blue (Sigma), calcium phosphate transfection reagents, peptide ligand (selected above), pCEP4 (Invitrogen), TRIZOL® Reagent, chloroform, isopropyl alcohol, 75% ethanol (in DEPC-treated water), 0.01%(v/v) DEPC water, SuperScripII reverse transcriptase (Life Technology), TE buffer (10 mM Tris (pH7.6), 1 mM EDTA), ethanol, 4M ammonium acetate (pH7.0), 1-kb DNA ladder, 10 x TBE electrophoresis buffer, 1% agarose in 1xTBE buffer with ethidium bromide.

Primers for RT-PCR are listed in Table 3.

Table 3. RT-PCR primer sequences.

Name	Sequence
RT-WAF1-FOR	5'-CTACCTCAGGCAGCTCAAGC-3'
RT-HME1-FOR	5'-AGACAGCACCCTCATCATGC-3'
RT-CD95-FOR	5'-TGGTGCTCATCTTAATGGCC-3'
RT-ACTIN-FOR	5'-TGACAAAACCTAACTTGCGC-3'
CDS Primer	5'-AAGCAGTGGTAACAACGCAGAGTACT$_{(30)}$N$_{-1}$N-3'
PCR Primer	5'-AAGCAGTGGTAACAACGCAGAGT-3'

Target genes and their corresponding RT-PCR product using the above primers are listed in Table 4.

Table 4. RT-PCR primer target genes and their corresponding product.

Primer name	Target gene		Amplify region in target mRNA	Size of RT-PCR product
	Name	Function		
RTWAF1FOR	p21	G1 arrest	1514-2121	607 bp
RTHME1FOR	14-3-3σ	G2 arrest	652-1245	593 bp
RTCD95FOR	CD95	Apoptosis	1968-2534	566 bp
RTACTINFOR	β-actin	cytoskeleton	1236-1793	557 bp

Cells SW480.7, U-2OS, and Sa-os2 were transfected with peptide ligand and pCEP4 parallelly using the calcium phosphate method. The transfection plan was shown in Table 5:

Table 5. Transfection plan for the verification of peptide ligand function.

Cell type	SW480.7		U-2OS		Sa-os2	
DNA	Ligand	pCEP 4	Ligand	pCEP 4	Ligand	pCEP4
Number of 6-well plates	3	3	1	1	1	1

After transfection, cells were cultured without drug selection and diluted at the ratio of 1:100 every five days. Cell proliferation status was closely observed until ligand transfected SW480.7 cells died out. Life staining with 0.4% trypan blue was applied to determine the death of the cells. Dead cells can not exclude the dye thus were stained blue. Living cells can pump the dye outside thus were not stained.

Total RNA was collected from two wells of one ligand transfected SW480.7 plate and one pCEP4 transfected SW480.7 plate each day on day 0, day 1, and day 2 after transfection. Thus RNA samples were harvested as SW480.7/Ligand-0day, SW480.7/Ligand-1day, SW480.7/Ligand-2day, SW480.7/pCEP4-0day, SW480.7/pCEP4-1day, SW480.7/pCEP4-2day. Protocol for total RNA extraction by TRIZOL® Reagent was according to that given by Life Technology:

1. Homogenization
   Wash the cell with DEPC water once, then add 1ml of TRIZOL Reagent per well. Pass the cell lysate several times through a pipette.
2. Phase separation
   Incubate the homogenized samples for 5 minutes at room temperature to permit the complete dissociation of nucleoprotein complexes. Add 200 μl of chloroform per 1 ml of TRIZOL Reagent. Cap the sample tubes securely. Shake tubes vigorously by hand for 15 seconds and incubate them at room temperature for 2 to 3 minutes.

Centrifuge the samples at no more than 12,000 x g for 15 minutes at 2 to 8°C.

3. RNA precipitation

Transfer the aqueous phase to a fresh tube. Precipitate the RNA by adding 500 µl of isopropyl alcohol per 1 ml of TRIZOL Reagent used for the initial homogenization. Incubate samples at room temperature for 10 minutes and centrifuge at no more than 12,000 x g for 10 minutes at 2 to 8°C.

4. RNA wash

Remove the supernatant. Wash the RNA pellet once with 75% ethanol, adding at least 1 ml of 75% ethanol per 1 ml of TRIZOL Reagent used for the initial homogenization. Mix the sample by vortexing and centrifuge at no more than 7,500 x g for 5 minutes at 2 to 8°C.

5. Re-dissolving the RNA

Briefly dry the RNA pellet. Dissolve RNA in DEPC water by passing the solution several times through a pipette tip, and incubation for 10 minutes at 55 to 60°C. Determine the concentration by measure OD_{260}. Store at -70°C.

<u>RT-PCR</u>:

First-Strand cDNA Synthesis

1. For each RNA sample, combine the following reagents in a sterile 0.2-ml reaction tube:

1-3 µl	RNA sample (using same amount of RNA among samples)
1 µl	CDS primer (10 µM)
1 µl	RT forward primer (10 µM)
x µl	Deionized H_2O
5 µl	Total volume

2. Mix and spin the tube briefly.
3. Incubate the tube at 70°C in a thermal cycler for 2 min.
4. Spin the tube briefly to collect contents at the

bottom. Keep tube at room temperature.

5. Add the following to each reaction tube:

2	μl	5 x First-Strand Buffer
1	μl	DTT (20 mM)
1	μl	50 x dNTP (10 mM)
1	μl	SuperScripII reverse transcriptase (200 units/μl)

6. Gently vortex and spin the tubes briefly.
7. Incubate the tubes at 42°C for 1 hr in an air incubator or cycler.
8. Add 40 μl TE buffer (10 mM Tris (pH7.6), 1 mM EDTA) to dilute the first-strand reaction product.
9. Heat tubes at 72°C for 7 min to inactivate reverse transcriptase.
10. Samples can be stored at -20°C for up to three months.

<u>cDNA Amplification by PCR</u>

1. Preheat the PCR thermal cycler to 95°C.
2. For each reaction, 5 μl of diluted first-strand cDNA and 5 μl of deionized water are added to a labeled 0.2-ml reaction tube.
3. Prepare a master mix for all reaction tubes, plus one additional tube. For each reaction:

20.5	μl	Deionized water
5	μl	10 x PCR buffer
1	μl	50 x dNTP (10 mM)
1.5	μl	PCR primer (10 μM)
1.5	μl	RT forward primer (10 μM)
10	μl	5x Q solution
0.5	μl	HotStart Taq DNA polymerase
40	μl	Total volume

4. Mix by vortexing and spin the tube briefly.
5. Aliquot 40 μl of the PCR Master Mix into each tube from Step 2.

6. Cap the tube, and place it in the preheated thermal cycler.

7. PCR was carried out using the following cycling parameters:

95 °C		15	min
2 cycles			
	94 °C	30	sec
	70 °C	30	sec
	72 °C	1	min
2 cycles			
	94 °C	30	sec
	69 °C	30	sec
	72 °C	1	min
2 cycles			
	94 °C	30	sec
	67 °C	30	sec
	72 °C	1	min
2 cycles			
	94 °C	30	sec
	65 °C	30	sec
	72 °C	1	min
2 cycles			
	94 °C	30	sec
	63 °C	30	sec
	72 °C	1	min
2 cycles			
	94 °C	30	sec
	61 °C	30	sec
	72 °C	1	min
2 cycles			
	94 °C	30	sec
	59 °C	30	sec
	72 °C	1	min
2 cycles			
	94 °C	30	sec
	57 °C	30	sec
	72 °C	1	min
40 cycles			
	94 °C	30	sec
	55 °C	30	sec
	72 °C	1	min
1 cycle			
	72 °C	7	min
4 °C	store		

8. Electrophoreses 5 μl of each PCR reaction alongside 0.1 μg of 1-kb DNA ladder on a 1.2% agarose/EtBr gel in 1 x TBE buffer.

Results

After transfecting the ligand into SW480.7 cells for 15 days without puromycin selection, the cells could no longer attach to the wall and died. On the contrary, the ligand-transfected U2-OS and Saos-2 cells were growing well.

RT-PCR analysis of the mRNA from ligand transfected SW480.7 cells showed that this ligand can reestablish wild type p53 function. Higher mRNA levels of p21 (WAF1/Cip1), 14-3-3σ and CD95 genes were found compared with that in the empty pCEP4 transfected SW480.7 cells. This proved that the apoptosis and cell growth arrest of ligand transfected cells were owing to the ligand that corrected R273H/P309S p53 mutant to wild type function.

__New reporter system__

The previous p53 reporter p21ur is based on the transactivation activity of wild type p53 to p21 promoter. It has been proved that up-regulation of p21 protein expression will result in cell growth arrest, but not apoptosis. Another p53 target, Bax, can lead to apoptosis once its transcription and expression is upregulated. Bax is one of the key mediators for p53 to induce cell apoptosis - to kill the cells that have failed to repair their damaged genomes.

Bax promoter is different from p21 promoter, though both have p53 consensus binding sequence. It has been reported that some p53 mutant can transactivate p21 promoter, but not Bax. This indicates a more stringent requirement of p53 conformation for the transactivation of Bax promoter.

Green fluorescence protein (GFP) has also been used in the new reporter system. Bax promoter drives a long

transcript of both puromycin resistant gene and GFP gene. By inserting an IRES fragment between the two genes that introduces a second ribosome-binding site, they can be translated into protein simultaneously from one transcript (Clontech pIRESneo manual). Thus the green color from GFP will visualize the Bax promoter activity, hence the ligand's activity to correct p53 mutants, in addition to the positive selection of puromycin resistance. The green fluorescence will eliminate the background caused by mechanisms which have cells developed to survive from puromycin selection other than transactivation of Bax promoter.

Selection marker is important for integration of the plasmid into genome. A stable p53-null cell line with integrated reporter system will be convenient for ligand screening, especially for non-peptide ligands.

Materials and Methods

Plasmids used in this Example are: p21ur (above), pIRESneo (Clontech), pEGFP-C1 (Clontech), and pREP8 (Invitrogen).

Genomic DNA (human, male, normal)

Restriction endonucleases (Promega): AgeI, EcoRV, HindIII, NruI, PvuI, SacI, SalI, SpeI, XbaI, XhoI, XmaI.

Modification enzymes: Vent DNA polymerase (NEB), T4 DNA ligase (Promega), and Shrimp alkaline phosphotase (SAP).

Bax promoter primer:

Bax promoter Forward:

5'-atctaagcttgaggcttcagcccgggaattccag-3'
HindIII

Bax promoter Reverse:

5'-atctaccggtgccagcagtggcgccgtccaacag-3'
AgeI

EGFP primer:
EGFP Forward: 5'-aataa <u>cccgGG</u>TCGCCACCATGGTGAGCAAG-3'
XmaI

EGFP Reverse: 5'-aataa <u>tctaga</u>ACTTGTACAGCTCGTCCATGCCG-3'
XbaI
stop

<u>Construction of pBaxur - Bax promoter controlled puromycin resistant gene</u>:

- Making PCR from normal human genomic DNA using Bax promoter primers. The PCR product is 500bp. It has a SmaI site near its 5'-end and a SacI site near its 3'-end.
- Cut PCR product Bax promoter with HindIII and AgeI.
- Cut p21ur with HindIII and AgeI, and dephosphorylate afterwards. Purify the 3,9kb fragment from gel.
- Ligate Bax promoter into the 3,9kb fragment so that Bax promoter can control the puromycin resistant gene expression in mammalian cells. The final product is denoted as pBaxur.

<u>Construction of pBaxurEGFP and p21urEGFP - Green fluorescent protein downstream of puromycin resistant gene that visualizes the transcription level of Bax or p21 promoter</u>:

- Amplify EGFP gene fragment from pEGFP-C1 vector by PCR with the EGFP primers.
- The PCR product is around 750bp. Cut the product with XmaI and XbaI, then purify it from gel.
- Cut pIRESneo with XmaI and XbaI, then purify the 4,5kb fragment from gel.
- Ligate PCR product with pIRESneo 4,5kb fragment to generate vector pIRES-EGFP (5,2kb).
- Cut p21ur or pBaxur with EcoRV and XbaI, and purify the 3.2kb or 1.3kb fragment from gel.
- Cut pIRES-EGFP with NruI and SpeI, dephosphorylate, and purify the 4.4kb fragment from gel.
- Ligate the two fragments to generate transient p53-double-reporter p21urEGFP (7.6kb) or pBaxurEGFP (5.7kb).

<u>Construction of pBaxurEGFP-hisD - Integrative plasmid with selection mark as histidinol D resistant</u>:

- Cut pREP8 with PvuI, SalI, and SacI, purify the 5194bp fragment from gel.
- Cut pBaxEGFP with PvuI and XhoI, dephosphorylated, and purify the 3972bp fragment from gel.
- Ligate the hisD fragment into pBaxurEGFP fragment to generate integrative p53-double reporter pBaxurEGFP_hisD.

The structure of p21urEGFP was verified by restriction of HindIII, which should give five bands of 3489bp, 2339bp, 1041bp, 360bp, and 348bp. Agarose gel electrophoresis of the restricted plasmid showed the expected bands.

The structure of pBaxur was verified by restriction of either HindIII together with SacI, which should give two bands of 3996bp, and 388bp, or SmaI alone, which should give three bands of 3184bp, 620bp, and 580bp. Agarose gel electrophoresis of the restricted plasmid showed the expected bands.

The structure of pBaxurEGFP was verified by restriction of either EcoRI together with XbaI, which should give two bands of 2907bp, and 2766bp, or SacI alone, which should give two bands of 3125bp, and 2548bp. Agarose gel electrophoresis of the restricted plasmid showed the expected bands.

The structure of pBaxurEGFP-hisD was verified by restriction of either EcoRI, which should give two bands of 5955bp, and 3211bp, or SacI, which should give two bands of 6618bp, and 2548bp. Agarose gel electrophoresis of the restricted plasmid showed the expected bands.

Screening for new peptide ligands is done by first transfecting the outlined new reporter p21urEGFP-hisD along with the cotransfection of the peptide library constructed above into cancer cell lines such as SW480.7. Caspase inhibitor VAD-fmk (ApoAlert®, Clontech) is added

to the cell culture at the final concentration of 40mM in order to prevent apoptosis in the case of the overexpression of BAX gene. Also 0.5μg/ml puromycin is added in the media one-day after transfection. After several days, surviving green cells (excited at 488nm wavelength) will be the cells containing the cadidate ligands. Rescued plasmids from these cells are subjected to DNA sequencing to deduce the amino acid sequences of the encoded peptide ligands.

When screening non-peptide ligands, a stable cell line of the reporter p21urEGFP-hisD is constructed. Members of the chemical compound library are added to the cell culture. The cells should also be cultured in the medium containing 40mM VAD-fmk and 0.5μg/ml puromycin. Surviving green cells after excitation at 488nm is the indicator for successfull candidate ligands.

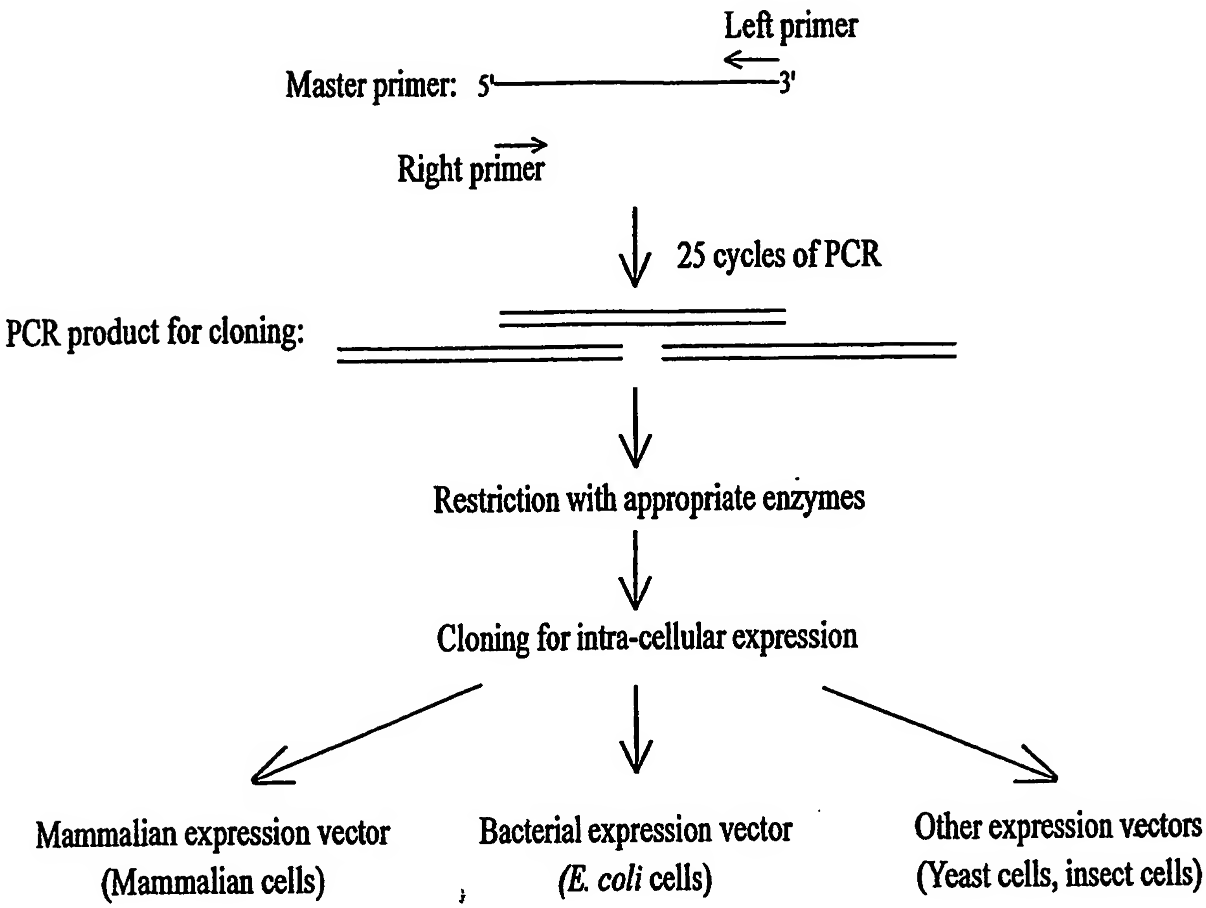

Left primer
Master primer: 5' 3'
Right primer
25 cycles of PCR
PCR product for cloning:
Restriction with appropriate enzymes
Cloning for intra-cellular expression
Mammalian expression vector
(Mammalian cells)
Bacterial expression vector
(*E. coli* cells)
Other expression vectors
(Yeast cells, insect cells)

Figure 1

Master primer:

5′-AA GAGCTC GGTACC AAG AAGGAG TTTACAT ATG GGA NNK NNK NNK TGA TAA
SacI KpnI SD Spacer M G X X X stop stop

GGATCC AAGCTT GAATTC AG-3′
BamHI HindIII EcoRI

Left primer:
5′-AA GAGCTC GGTACC AAG AAGGAG-3′
SacI KpnI SD

Right primer:
5′-CT GAATTC AAGCTT GGATCC TTATC-3′
EcoRI HindIII BamHI

Figure 2

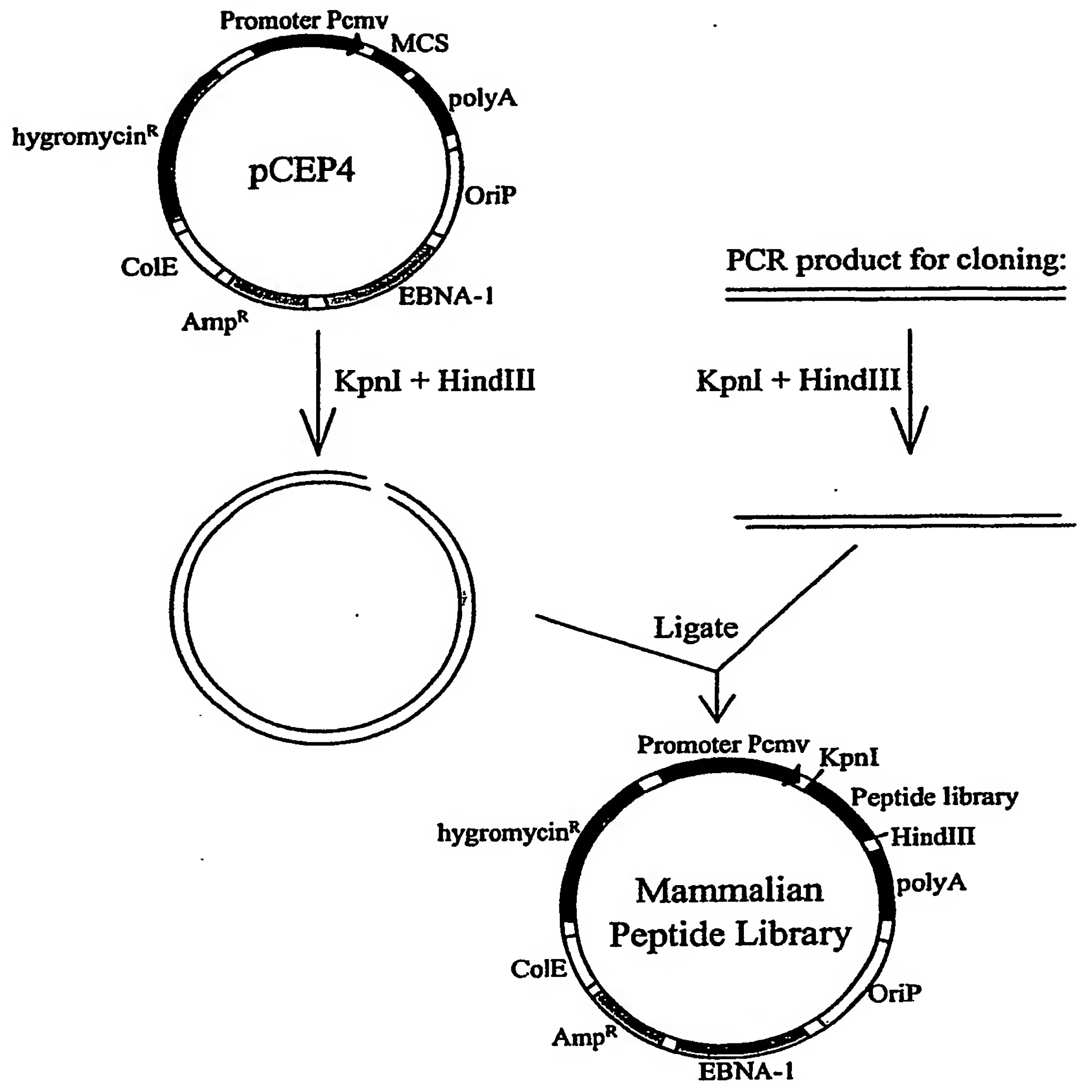

Promoter Pcmv
MCS
polyA
hygromycin^R
pCEP4
OriP
ColE
EBNA-1
Amp^R
PCR product for cloning:
KpnI + HindIII
KpnI + HindIII
Ligate
Promoter Pcmv
KpnI
Peptide library
HindIII
hygromycin^R
polyA
Mammalian
Peptide Library
ColE
OriP
Amp^R
EBNA-1

Figure 3

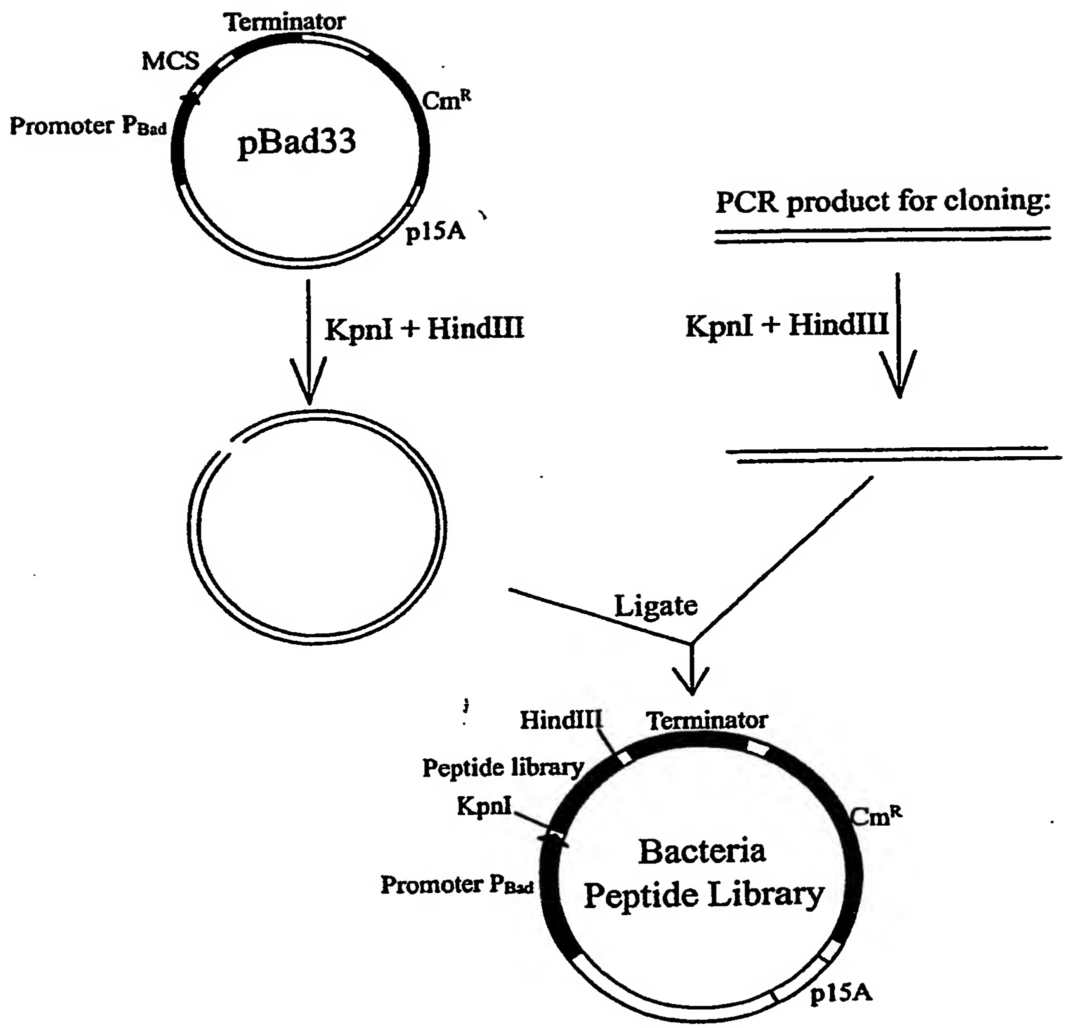

Terminator
MCS
Promoter PBad
pBad33
CmR
p15A
PCR product for cloning:
KpnI + HindIII
KpnI + HindIII
Ligate
HindIII
Terminator
Peptide library
KpnI
CmR
Promoter PBad
Bacteria
Peptide Library
p15A

Figure 4

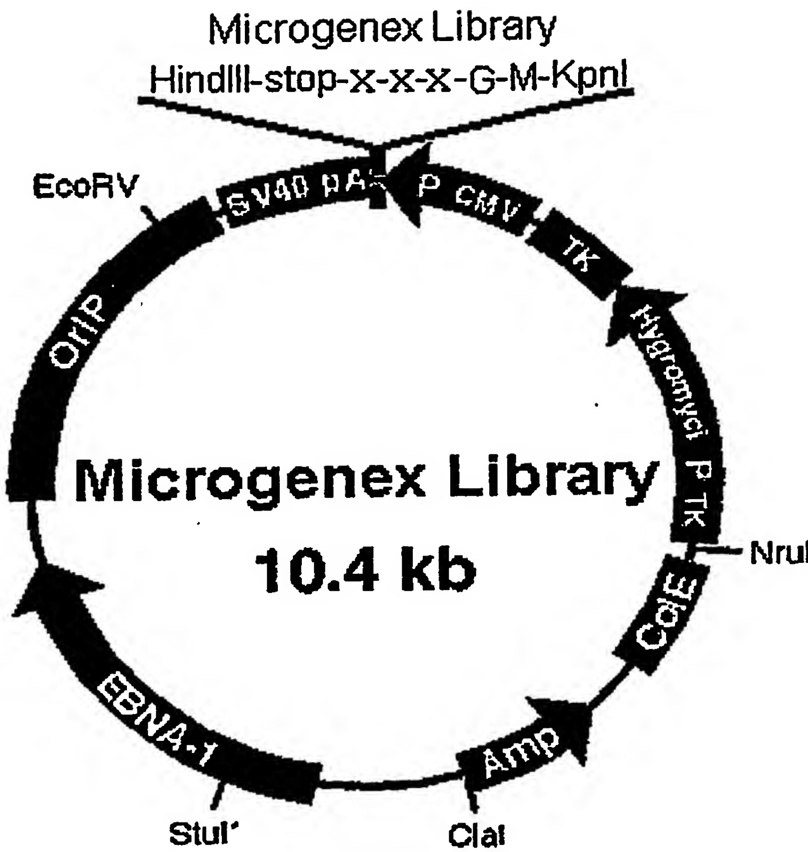
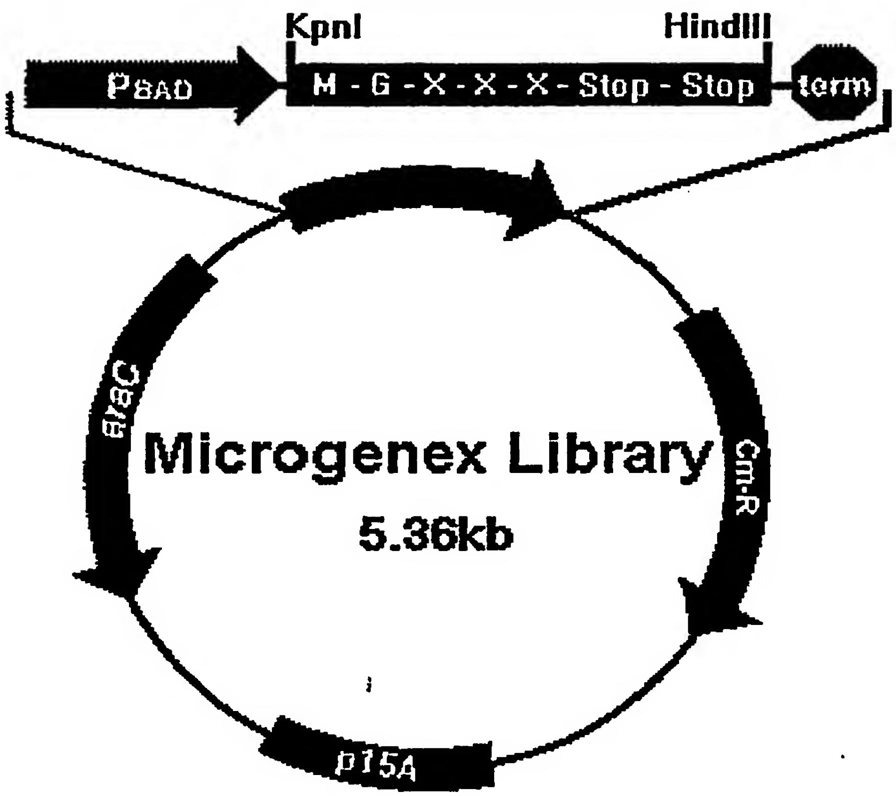

Microgenex Library
HindIII-stop-x-x-x-G-M-KpnI
EcoRV
SV40 pA
P CMV
TK
Hygromyci
P TK
NruI
ColE
Amp
ClaI
StuI'
EBNA-1
OriP
Microgenex Library
10.4 kb

Figure 5



KpnI
HindIII
PBAD
M - G - X - X - X - Stop - Stop
term
araC
Cm-R
p15A
Microgenex Library
5.36kb

Figure 6

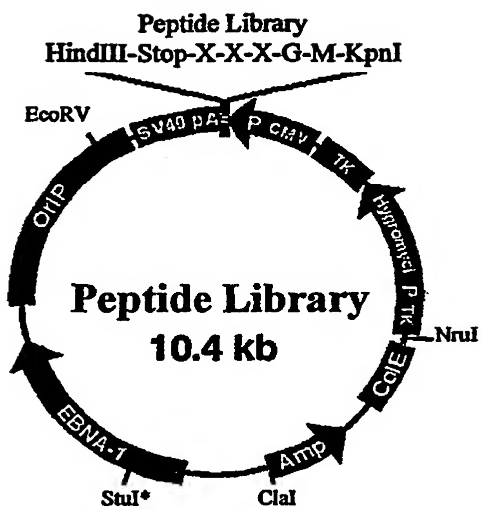

Peptide Library
HindIII-Stop-X-X-X-G-M-KpnI
EcoRV
SV40 pA
P CMV
TK
Hygromycin
P TK
NruI
ColE
Amp
ClaI
StuI*
EBNA-1
OriP
Peptide Library
10.4 kb

Figure 7